

УДК 004.051

А.С. Шаповалова

Тернопільський національний технічний університет імені Івана Пулюя, Україна

РОЗРОБКА АРХІТЕКТУРИ МЕТА-ПОШУКОВОЇ СИСТЕМИ "ЇЖАЧОК"

A.S. Shapovalova

DEVELOPMENT OF THE ARCHITECTURE OF THE META-SEARCH ENGINE "HEDGEHOG"

При дослідженні існуючих алгоритмів та методів пошуку в сучасних пошукових системах за показниками якості для виявлення основних факторів, що впливають на ранжирування сайтів в результатах пошуку було виявлено проблеми, пов'язані з пошуковим спамом і пошуком інформації в мережі в цілому. На основі результатів дослідження було запропоновано та розроблено власну мета-пошукову систему під назвою «Їжачок», яка дозволила об'єднати лідерів експериментів з пошуку інформації в інтернет павутині.

Для досягнення поставленої мети було виконано такі етапи роботи:

- проведено аналіз сучасних пошукових систем, використовуваних в українському та світовому сегменті мережі Інтернет;
- розроблено систему факторів, що беруть участь у формулах ранжування пошукових систем;
- здійснено огляд алгоритмів пошукових систем;
- розглянуто теоретичні підходи до обґрунтування проблеми пошукової оптимізації та ранжування web-сайтів;
- з'ясовано сутність пошукової оптимізації та ранжирування як предметів дослідження, охарактеризовано чинники ранжування;
- досліджено специфіку роботи сучасних пошукових машин;
- надано рекомендації щодо розробки програмного модуля на основі власного алгоритму пошукової системи.

В якості досліджуваних пошукових систем, для яких проводилася оптимізація, були обрані найбільш популярні та прогресивні Яндекс, Yahoo та Faroo.

Для вирішення поставлених завдань використані методи теорії множин, теорії систем масового обслуговування, порівняльного аналізу, об'єктно-орієнтованого аналізу, розрахована спрощена формула для визначення релевантності сторінки сайту з врахуванням сучасних механізмів, за допомогою схем, графіків, діаграм і спостережень.

Практична цінність роботи полягає в можливості використання отриманих науково-технічних результатів при експлуатації, дослідженні, що вимагають відносного порівняння альтернативних алгоритмів і методів.

Зокрема одним з етапів реалізації розробки проекту було створити трьохрівневу архітектуру мета-пошукової системи «Їжачок» (рис.1) та описано функціонал на кожному з рівнів.

Data access layer. Цей рівень забезпечує доступ до даних. Він інкапсулює арі пошукових систем, але замість запиту до бази даних, він робить http-запит до арі пошукової системи, і повертає результат пошуку на рівень виклику.

Logic Layer. Цей рівень містить в собі всю цінність системи та засоби для поєднання всіх пошукових систем. Він містить json-файли з конфігурацією пошукових систем, ключі доступу, аналіз та обробку даних, методи зведення результатів різних

пошукових систем до спільного інтерфейсу, та сервіси для зручнішої взаємодії з data access layer.

Presentation Layer. Цей рівень є рівнем представлення даних та взаємодією з користувачем. Він містить в собі контролери обробки http-запитів користувача, повернення та формування html-сторінок з графічним інтерфейсом, та результатами пошуку. Також цей рівень є основним в системі, так як містить в собі всі налаштування веб-серверу, і здійснює взаємодію з браузером (а це означає що він напряду взаємодіє з користувачем).

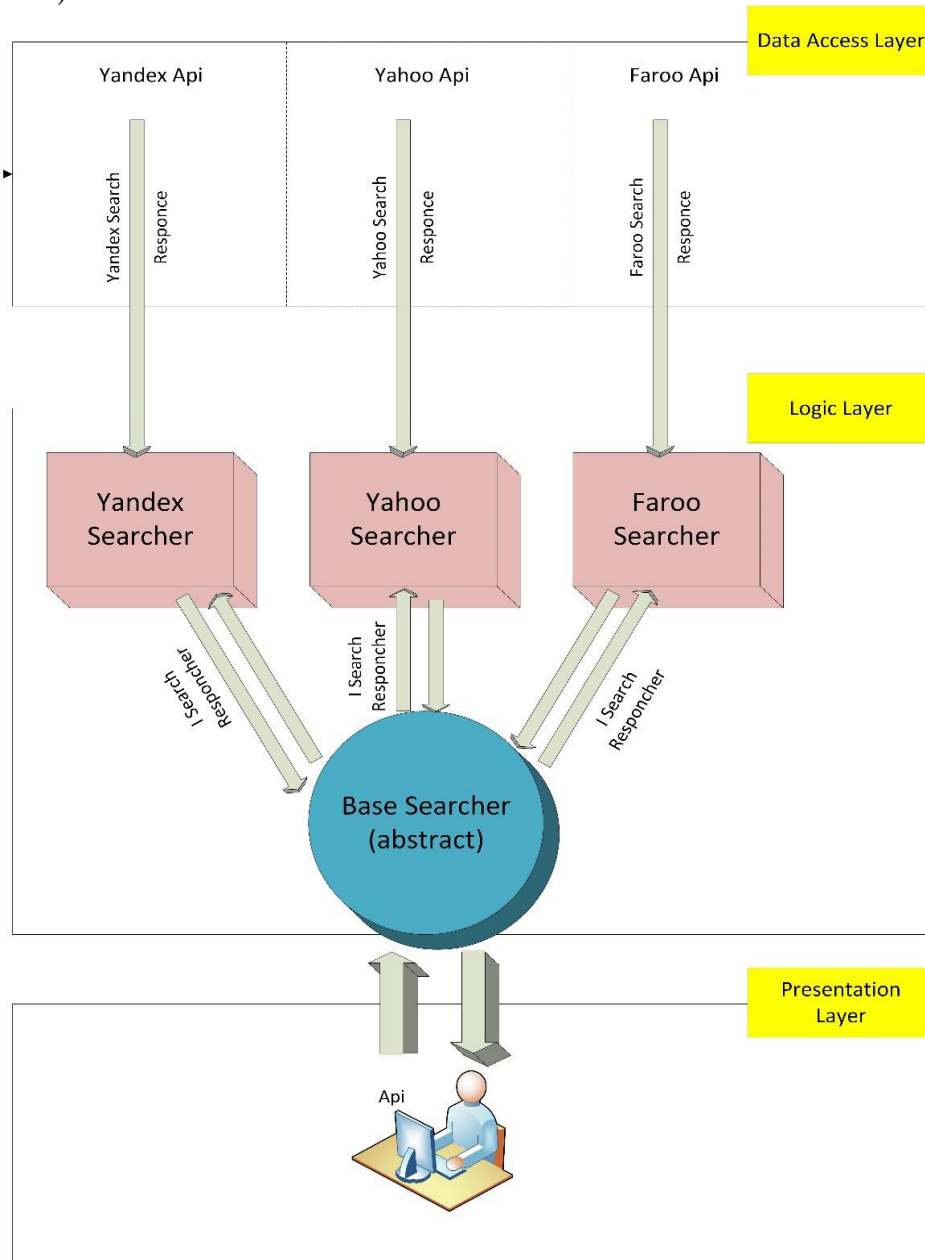


Рисунок 1. Архітектура пошукової системи «Їжачок»